Supporting Information

Genetic incorporation of L-dihydroxyphenylalanine (DOPA) biosynthesized by a tyrosine phenol-lyase

Sanggil Kim[1], Bong Hyun Sung[2], Sun Chang Kim[3], and Hyun Soo Lee[1],*

[1] *Department of Chemistry, Sogang University, Seoul 121-742, Republic of Korea.*

[2] *Bioenergy and Biochemical Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon 305-806, Republic of Korea*

[3] *Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Republic of Korea*

■ List of Contents

Materials and methods

General

All chemicals and DNA oligomers were obtained from commercial sources and used without further purification. Protein MS analysis was carried out using a Bruker Autoflex Speed MALDI-TOF mass spectrometer (Bruker Daltonics, Leipzig, Germany). All fluorescence spectra were obtained using the Hitachi F-7000 fluorescence spectrophotometer.

Screening aminoacyl-tRNA synthetase mutants

A mutant gene library of Methanococcus jannaschii TyrRS was synthesized by overlap extension PCR using pBK-DOPA-RS[1] as a template, and DNA primers with randomized sequences (NNK) at Asp158 and Ala167. The PCR product was digested with PstI and EcoRI, and ligated into a pBK vector digested with the same restriction enzymes. The plasmid DNA containing the mutant gene library was transformed into DH10β electrocompetent cells containing pREP[2], which harbors a chloramphenicol acetyltransferase (CAT) gene with an amber mutation at position 112. Transformants were cultivated in lysogeny broth (LB) medium containing kanamycin and tetracycline, and cells were harvested after 12 hours of cultivation. The cells (2 mL) were transferred to a glycerol minimal medium[2] (100 mL) containing kanamycin (50 μg/mL), tetracycline (12.5 μg/mL), 100 μM DTT, 1 mM DOPA and chloramphenicol (35 μg/mL), and incubated at 37 °C for 16 hours. Surviving cells were transferred to an LB agar plate containing kanamycin and tetracycline, and pBK plasmids from the colonies were isolated and characterized by sequencing**.**

Expression and purification of emGFP mutants containing DOPA

The emGFP gene was obtained from a commercial source by gene synthesis, amplified by PCR, and inserted between the BspHI and KpnI sites of pBAD/Myc-His (Invitrogen) to generate pBAD-emGFP. An amber codon (TAG) was introduced at position 39 or 90 in emGFP by site-directed mutagenesis. The plasmid containing each amber mutation was co-transformed with pEvol[3]-DOPA-RS, pEvol-DOPA-RS1 or pEvol-DOPA-RS2 into *E. coli* DH10β. Cells were then cultured in LB medium supplemented with ampicillin (100 μg/mL) and chloramphenicol (35 μg/mL). The starter culture (2 mL) was transferred to a defined medium[4] (100 mL) (50 mM Na_2HPO_4, 50 mM KH_2PO_4, 25 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% (w/v) trace metals, 0.5% (w/v) glycerol, 0.05% (w/v) glucose, and 0.36% (w/v) amino acids) supplemented with ampicillin (100 μg/mL), chloramphenicol (35 μg/mL), 300 μM DTT, and 3 mM DOPA at 30 °C. Protein expression was induced by adding 0.2% (w/v) L-arabinose when the culture optical density reached 0.8, and the culture was grown overnight at 30 °C. Cells were harvested by centrifugation, resuspended in a lysis buffer (10 ml) (50 mM Na_2HPO_4, 300 mM NaCl, 10 mM imidazole, and pH 8.0) and sonicated. Target proteins were purified by Ni-NTA affinity chromatography under native conditions according to the manufacturer's instructions (Qiagen). Protein concentrations were calculated by measuring absorbance at 280 nm and using the calculated extinction coefficient (2.2×10^4 $cm^{-1}M^{-1}$ for emGFP) (http://www.biomol.net/en/tools/proteinextinction.htm).

Fluorescence measurement of emGFP

Cells (4×10^9) were harvested by centrifugation at 10000 rpm and 4 °C for 5 minutes, and cell pellets were lysed for 1 hour with Bugbuster (Novagen) (100 μL) supplemented with a benzonase nuclease (Sigma) (250 units/μL). Cell debris was removed by centrifugation at 13000 rpm and 4 °C for 10

minutes, and the supernatant was directly used for fluorescence measurement. Fluorescence was measured at 510 nm with excitation at 487 nm.

Genetic incorporation of DOPA biosynthesized from catechol, pyruvate, and ammonia

The TPL gene was amplified from *Citrobacter freundii* genomic DNA (ATCC8090) and inserted between the NcoI and PstI sites of a pHCE vector (Takara Bio. Inc.) to generate pHCE-TPL. The TPL gene, including the flanked promoter and terminator sequences in pHCE-TPL, was amplified and inserted into the BsrGI site of pBAD/Myc-His (Invitrogen) to produce pBAD-dual-TPL. The C-terminal hexahistidine-tagged emGFP-E90TAG gene was amplified from pBAD-emGFP-E90TAG and inserted into the NcoI and KpnI sites of pBAD-dual-TPL to generate pBAD-dual-TPL-emGFP-E90TAG. The plasmid was co-transformed with pEvol-DOPA-RS2 into *E. coli* DH10β, and the transformed cells were amplified in LB medium supplemented with ampicillin (100 μg/mL) and chloramphenicol (35 μg/mL). The starter culture (2 mL) was transferred to a defined medium[4] (100 mL) (50 mM Na_2HPO_4, 50 mM KH_2PO_4, 25 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% (w/v) trace metals, 0.5% (w/v) glycerol, 0.05% (w/v) glucose, and 0.36% (w/v) amino acids, and pH 7.25) supplemented with ampicillin (100 μg/mL), chloramphenicol (35 μg/mL), 100 mM pyruvate, 10 mM catechol, 25 mM $(NH_4)_2SO_4$, and 300 μM DTT at 30 °C. Protein expression was induced by adding 0.2% (w/v) L-arabinose when the optical density reached 0.8, and the culture was incubated overnight at 30 °C. Cells were harvested by centrifugation, resuspended in a lysis buffer (10 ml) (50 mM Na_2HPO_4, 300 mM NaCl, 10 mM imidazole and pH 8.0), and sonicated. Target proteins were purified by Ni-NTA affinity chromatography under native conditions according to the manufacturer's protocol (Qiagen). Protein concentrations were calculated by measuring absorbance at 280 nm and using the calculated extinction coefficient (2.2×10^4 $cm^{-1}M^{-1}$ for emGFP) (http://www.biomol.net/en/tools/proteinextinction.htm).

SPOCQ cycloaddition reactions with a MBP mutant containing DOPA

Each stock solutions of MBP-K313DOPA (100 μM), sodium periodate (1 mM) and Cy5.5-ADIBO (5 mM) in 50 mM phosphate buffer (pH 8.0, prepared by mixing 3.71 mg K_2HPO_4 and 410 mg K_2HPO_4 in water to make 1.0 L buffer solution) containing 300 mM NaCl were mixed in the indicated final concentration (10 μL total volume). MBP-K313DOPA (20 μM, final concentration) was treated with sodium periodate (20 μM, final concentration) and Cy5.5-ADIBO (200 μM, final concentration), and the mixture was incubated at 25 °C. The reaction was quenched by adding excess L-dopaquinone at various time points (1, 5, 10, 30, and 60 min), and the mixture was analyzed by SDS-PAGE. Fluorescence images were obtained using a G:BOX Chemi Fluorescent & Chemiluminescent Imaging System (Syngene), and the gel was stained with Coomassie Brilliant Blue R-250.

Oligomerization of a MBP mutant containing DOPA

Each stock solutions of MBP-K83DOPA (400 μM) and sodium periodate (1 mM) in 50 mM phosphate buffer (pH 8.0, prepared by mixing 3.71 mg K_2HPO_4 and 410 mg K_2HPO_4 in water to make 1.0 L buffer solution) containing 300 mM NaCl were mixed in the indicated final concentration (10 μL total volume). MBP-K83DOPA (300 μM, final concentration) was treated with sodium periodate (300 μM, final concentration) for 48 hours at room temperature, and the mixture was analyzed by SDS-PAGE.

Protein-protein cross-linking with an affibody and Z-domain protein

Each stock solutions of AFB-D36DOPA (400 μM), sodium periodate (1 mM) and MBP-Z-N6K (20 μM) in 50 mM phosphate buffer (pH 8.0, prepared by mixing 3.71 mg K_2HPO_4 and 410 mg K_2HPO_4 in water to make 1.0 L buffer solution) containing 300 mM NaCl were mixed in the indicated final concentration (10 μL total volume). AFB-D36DOPA (200 μM, final concentration) was treated with sodium periodate (200 μM, final concentration) and MBP-Z-N6K (50 μM, final concentration), and the mixture was incubated for 4 hours at 25 °C. The reaction mixture was then analyzed by SDS-PAGE.

MALDI-TOF MS analysis

Protein (200 μg, 0.5 mg/mL final concentration) was digested with trypsin (20 μM) in a reaction buffer containing 50 mM Tris and SDS (10 mg/mL) at 37 °C for 12 hours, and the digested peptides were desalted using a C-18 spin column. The desalted tryptic peptides were mixed with α-cyano-4-hydroxycinnamic acid (CHCA) matrix (10 mg/mL in water containing 50% (v/v) acetonitrile and 0.1% (v/v) trifluoroacetic acid) at a 1:1 ratio (v/v) and subjected to MS analysis.

References and Notes

[1] L. Alfonta, Z. Zhang, S. Uryu, J. A. Loo and P. G. Schultz, J. Am. *Chem. Soc*., 2003, **125**, 14662.

[2] L. Wang, A. Brock, B. Herberich and P. G. Schultz, *Science*, 2001, **292**, 498.

[3] T. S. Young, I. Ahmad, J. A. Yin and P. G. Schultz, *J. Mol. Biol*., 2010, **395**, 361.

[4] F. W. Studier, *Protein Expres. Purif*., 2005, **41**, 207.

Figure S1. (A) Location of Y39 and E90 in GFP (left; front view, right; top view). These solvent-exposed residues were chosen to minimize structural perturbation when they were replaced with DOPA. (B) SDS-PAGE analyses of purified emGFP-Y39DOPA and emGFP-E90DOPA. The mutant GFPs were expressed in the presence of one of the evolved aaRS mutants and varied concentrations of L-DOPA, and purified by Ni-NTA affinity chromatography. The purified proteins were analyzed in the same amount (5 μg) and visualized by Coomassie-staining.

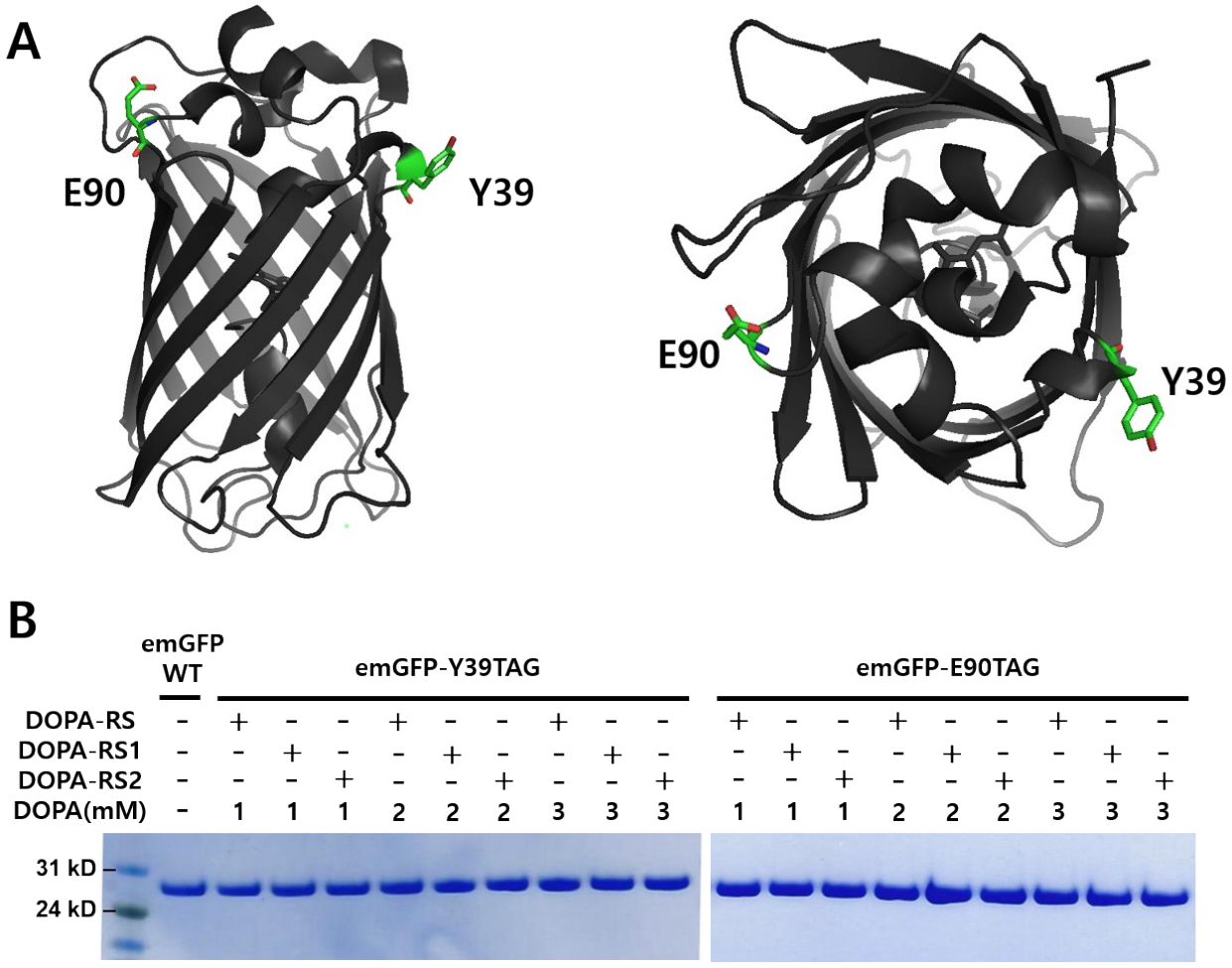

Figure S2. (A) Fluorescence measurements from emGFP mutants expressed using three aaRS mutants evolved for DOPA. DOPA was incorporated into two positions of emGFP: position 39 or position 90. DOPA-RS is the previously reported aaRS mutant for comparison, and DOPA-RS1 and DOPA-RS2 were evolved in the present study. n = 3, error bars; mean ± s.d. (B) MALDI-TOF MS analysis of emGFP-E90DOPA expressed by three aaRS mutants with 1, 2, and 3 mM DOPA. Peptide E (residues 86–96; SAMPEGYVQER) represents the tryptic peptide fragment containing E90. Peptide Y and Peptide DOPA represent the peptides containing Tyr and DOPA at position 90.

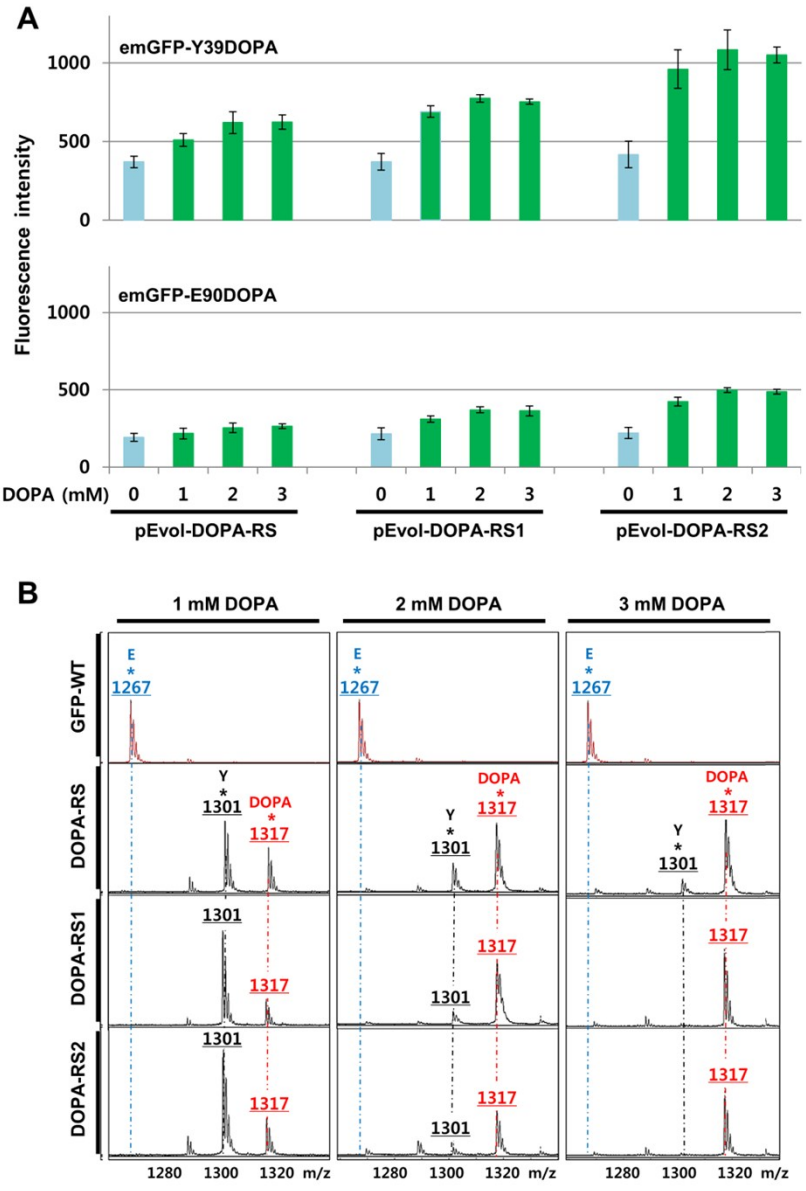

Figure S3. (A) SDS-PAGE analysis of whole cell lysates from cells expressing emGFP-E90DOPA by the designed biosynthetic system. Whole cell lysates from cells expressing emGFP-E90DOPA in the presence of DOPA were also analyzed as a control. Purified TPL-WT and emGFP-WT were analyzed for comparison. (B) SDS-PAGE analysis of purified emGFP-E90DOPA expressed by the designed biosynthetic system in the presence of 6, 8, or 10 mM catechol. The proteins were purified by Ni-NTA affinity chromatography, and the protein bands were visualized by Coomassie-staining.

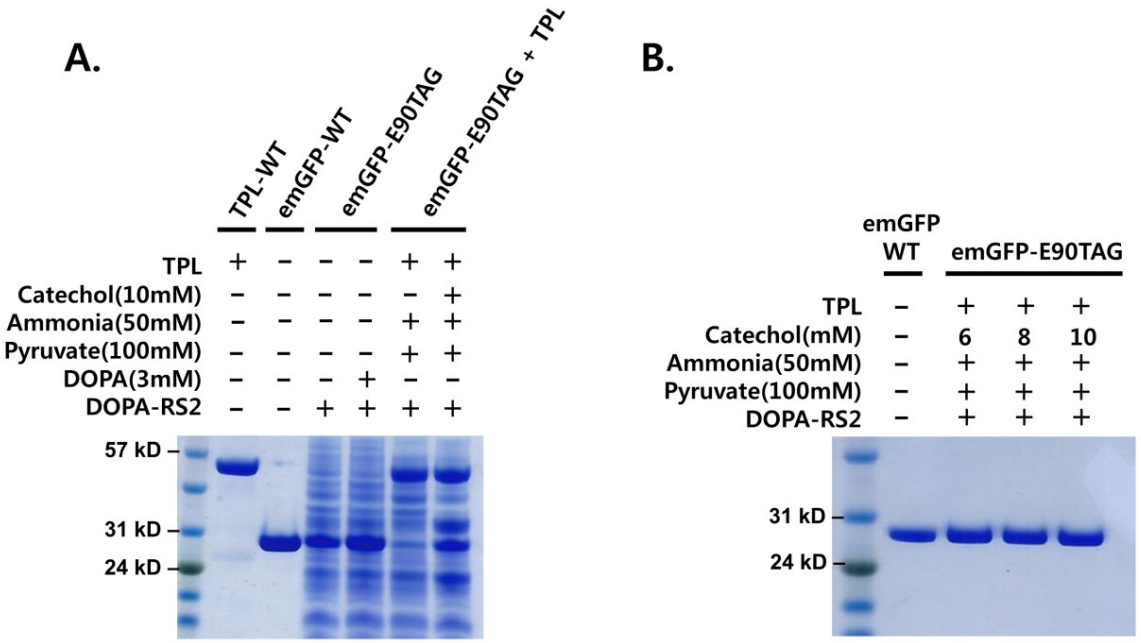

Figure S4. Comparison of cell density obtained by direct incorporation of DOPA biosynthesized from TPL with those of the general genetic incorporation of DOPA. n = 3, error bars; mean ± s.d.

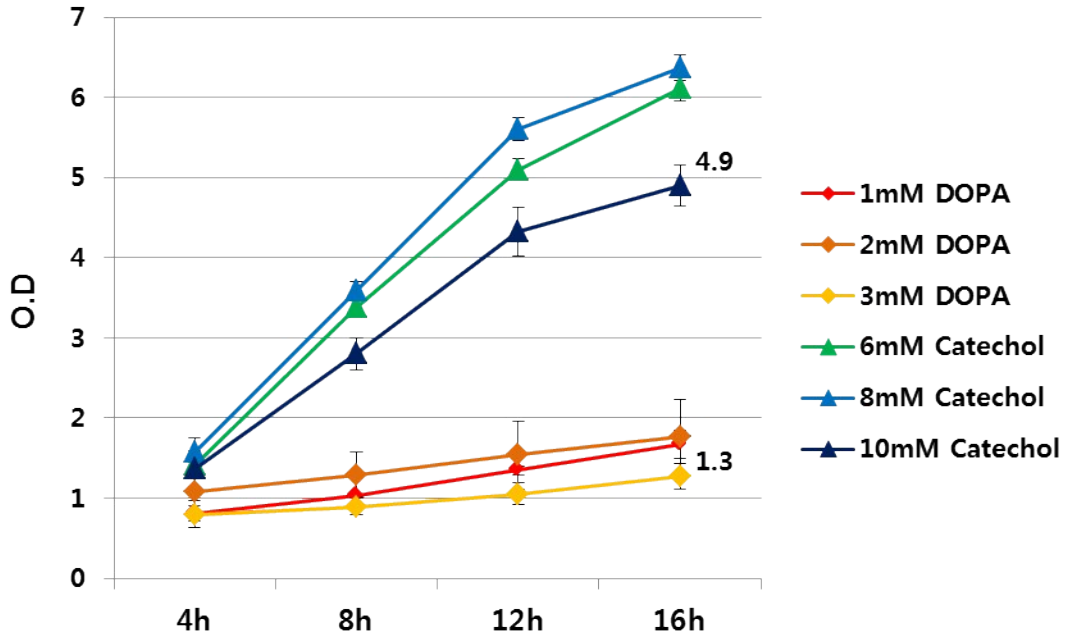

Figure S5. The reaction scheme for SPOCQ cycloaddition. DOPA is genetically incorporated into a protein and oxidized to dopaquinone, which is subject to SPOCQ cycloaddition to react with a strained alkyne.

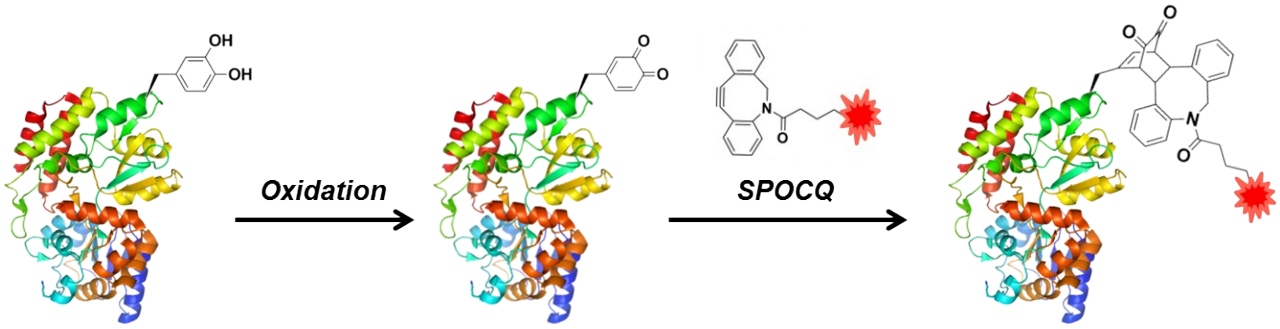

Figure S6. SPOCQ cycloaddition reactions of MBP-K313DOPA with ADIBO-Cy5.5. MBP-WT and MBP-K313DOPA were reacted with ADIBO-Cy5.5 in the presence (+) and absence (-) of $NaIO_4$ at room temperature for 60 min. The reaction mixtures were analyzed by SDS-PAGE, and the gel was visualized by Coomassie-staining and fluorescence.

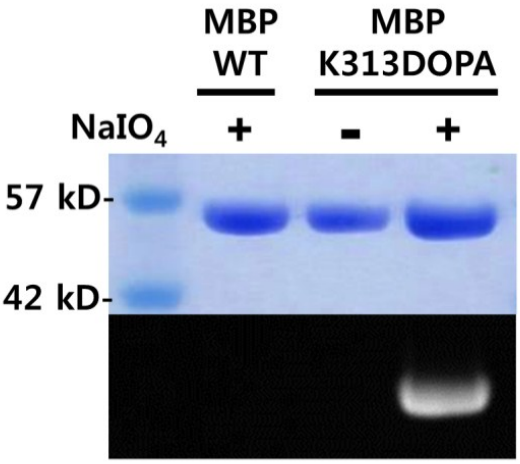

Figure S7. Control experiments for oligomerization of MBP-K83DOPA. MBP-WT and MBP-K83DOPA were incubated in the presence (+) and absence (-) of $NaIO_4$ at room temperature for the indicated time and analyzed by SDS-PAGE.

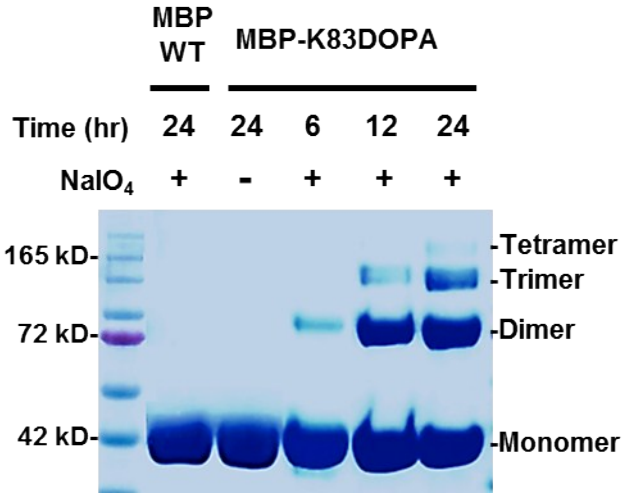

Table S1. Estimated cost for production of DOPA-containing proteins by the biosynthetic method in comparison with that by the general method using DOPA. Expanses for pyruvate and ammonium chloride used in the DOPA biosynthesis are negligible and not included in this calculation.

	Catechol (biosynthesis)	DOPA
Commercial price[a] ($/100 g)	21.6	448
Culture concentration (mM)	10	3
Protein yield[b] (mg/L)	5.09	1.8
Calculated expense for 1.0 mg protein production ($)	0.05	1.47

[a] averaged from prices in three major chemical companies.
[b] data from Table 1.